PINT: Pathways INtegration Tool

Y.-T. Wang[1], Y.-H. Huang[2], Y.-C. Chen[3], C.-L. Hsu[4] and U.-C. Yang[3,4,*]

[1]Institute of Biochemistry and Molecular Biology, College of Life Science, [2]Department of Biochemistry, Faculty of Medicine, School of Medicine, [3]Center for Systems and Synthetic Biology and [4]Institute of Biomedical Informatics, College of Life Science, National Yang-Ming University, No. 155, Sec. 2, Li-Nong St, Taipei City, Taiwan 11221, R.O.C

ABSTRACT

New pathway databases generally display pathways by retrieving information from a database dynamically. Some of them even provide their pathways in SBML or other exchangeable formats. Integrating these models is a challenging work, because these models were not built in the same way. Pathways integration Tool (PINT) may integrate the standard SBML files. Since these files may be obtained from different sources, any inconsistency in component names can be revised by using an annotation editor upon uploading a pathway model. This integration function greatly simplifies the building of a complex model from small models. To get new users started, about 190 curated public models of human pathways were collected by PINT. Relevant models can be selected and sent to the workbench by using a user-friendly query interface, which also accepts a gene list derived from high-throughput experiments. The models on the workbench, from either a public or a private source, can be integrated and painted. The painting function is useful for highlighting important genes or even their expression level on a merged pathway diagram, so that the biological significance can be revealed. This tool is freely available at http://csb2.ym.edu.tw/pint/.

INTRODUCTION

Public-domain pathway models have become an invaluable resource to the research community. Hundreds of expert-curated biological pathways are available in the public domain databases (1–3). Unlike conventional pathways that were purely static diagrams, many of these pathways have been prepared as text-based models in machine-readable formats such as Systems Biology Markup Language (SBML) (4) and Biological Pathways eXchange (BioPax) (http://www.biopax.org). These formats have been tailored to facilitate data storage and exchange as well as information retrieval, thus rendering high accessibility to pathway models. A number of pathway tools have therefore been designed to take advantage of pathway models to perform pathway analysis such as topological network analysis and mathematical simulation (5). Besides, bench biologists may use visualization tools to map data on top of the diagrams of such models to aid the interpretation of experiment results (6–8).

One issue concerning the application of public-domain pathway models is that each such model usually contains just partial information about the biological networks, covering only a subset of molecules that are implicated in a cellular response to stimuli. A plausible situation that biologists may encounter is that phenotypically relevant entities screened from high-throughput experiments may disperse across distinct pathway models. This means that, when investigating how such entities contribute to a phenotype, biologists may have to jump back and forth among several pathway diagrams to look for reactions participated by these entities. This process demands considerable effort and can be error-prone. We argue that one approach to improve this situation is to integrate multiple pathway models together to build a more comprehensive one. In the following text, we refer to this operation as biological pathways integration (BPI).

BPI is more than a trick to build bigger models. New models created through BPI may benefit biologists on several occasions. For instance, novel protein–protein interactions may suggest crosstalk among initially distinct pathway models (9); fusion genes in cancer cells may result in aberrant function links among normally independent pathways (10,11); promiscuous domains in certain cancer proteins may mediate abnormal protein–protein interactions (12), possibly leading to pathological pathway crosstalk; a therapeutic agent targeting against one malignant pathway might have failed due to yet

unconfirmed crosstalk from other pathways to compensate the blocked function (13). In either case mentioned above, a merged model of potentially associated pathways may provide a better overview of all implicated and suspicious entities—the first step toward interpreting biological observations.

Currently, BPI is not really supported by most public-domain pathway editors. Most such tools only permit entities to be added/deleted one by one, namely in an incremental manner. Albeit a general 'copy and paste' operation may place the topological structures of more than one pathway together, it cannot rebuild the relations among distinct models. SemanticSBML is the only tool we are aware that supports pathways integration. However, this tool appears to have been designed for experienced curators who are familiar with a number of entity annotation systems (14). Besides, installing the standalone version of this tool in a local machine is very difficult. Here, we present Pathways Integration Tool, PINT, a website that is free and open to all users and there is no login and installation requirement, to facilitate BPI. Our website was designed to make BPI a less painful task, not assuming that new users must know where to find useful pathway models to start an integration, what the difference is between various notation conventions, and how to find the UniProt or ChEBI accession numbers for entities. Besides, the PINT painting function can highlight particular entities on a merged diagram. We describe in the following sections about the implementation and the unique features of PINT. We also show an application of PINT to explore the features of cancer related pathways.

IMPLEMENTATION

The main subject of BPI is about how multiple pathway models can be integrated to generate a merged model, in which not only the original pathway information in each model is preserved but also the initially unavailable inter-pathway relations can be rebuilt. Hence, PINT has been implemented with a set of functions to assist users to finish the following steps: (i) users upload their pathway models into a BPI system; (ii) the BPI system pool users' pathway models together; (iii) the BPI system finds pairwise relation among pathway models; (iv) the BPI system rebuilds the inter-model connections; and (v) the BPI system removes the redundant nodes and edges from the merged model. In particular, PINT provides two useful features, a tolerant integration mode and an online annotation editor, to help manage the exceptions caused by the inter-model incompatibilities of the annotation styles and data storage formats, respectively. The PINT rule and core functions used to integrate pathways are further described in the following sections.

The pathways integration rule

In the process of integrating pathway models, the first step, perhaps the most important one, is to find the 'relations' among the models to be integrated. In the PINT BPI function, the concurrent occurrences of an entity across distinct models are considered as an inter-pathway relation. PINT takes such entities to be a putative 'linker' that can re-connect two initially distinct models.

PINT integrates multiple pathway models into a merged one through a process involving identification of candidate linkers among models and reconstruction of relations between initially distinct models (Figure 1A and B). Technically, PINT integrates pathways at the level of nodes, which may consist of one or multiple entities (i.e. a complex). To ensure that a merged model can convey accurate biological information, PINT employs a stringent rule to perform BPI. For integrating pathways *P1* and *P2*, as in Figure 1A, PINT uses the following criteria to determine if nodes *N1* and *N2* can be linkers:

(i) Node *N1* is in the same compartment as node *N2*.
(ii) Node *N1* consists of exactly the same set of entities as node *N2* does.
(iii) Entities considered to be identical are annotated with exactly the same properties (e.g. symbols, accession numbers and modifications).

Thus, nodes *N1* and *N2* as a whole is a putatively identical node (PIN) set; nodes *N3* and *N4* as a whole is another PIN set. Likewise, putatively identical edge (PIE) set is determined in a similar manner. Edges *X1* and *X2* as a whole is a PIE set.

To generate a merged model containing non-redundant information, PINT preserves only one node/edge from each PIN/PIE set when integrating models (Figure 1B, deleted nodes and edges are indicated by red crosses). Finally, to reconstruct inter-pathway links, PINT recursively relocates the end of each edge that was connected to each discarded node, to respectively identical node preserved in a merged model (see Figure 1B, edge *X4* is replaced by edge *X5*). Thus, initially unlinked pathway models can be re-connected via relocated edges.

Handling pathways-integration exceptions

In the process of BPI, various types of exceptions can occur, which may or may not be handled by PINT. In the following subsections, we show three situations that may cause exceptions and describe how PINT manages them.

Integrating pathways containing edges of inconsistent annotation of directionality. It is possible that in one model of a metabolic pathway a reaction is annotated as reversible (i.e. bi-directional edge), whereas in other models the counterpart reaction is annotated as irreversible (i.e. one-directional edge). If only a stringent rule of pathways integration were applied, such annotation inconsistency can lead to redundant edges between nodes in the merged model. Thus, PINT has been implemented with two additional rules to handle reaction directionality: (i) if an edge in one model is one-directional and its counterpart in another model is bi-directional, then in the merged model this edge will be assigned as bi-directional; (ii) if an edge in one model is one-directional and its counterpart in another model is also one-directional but in the
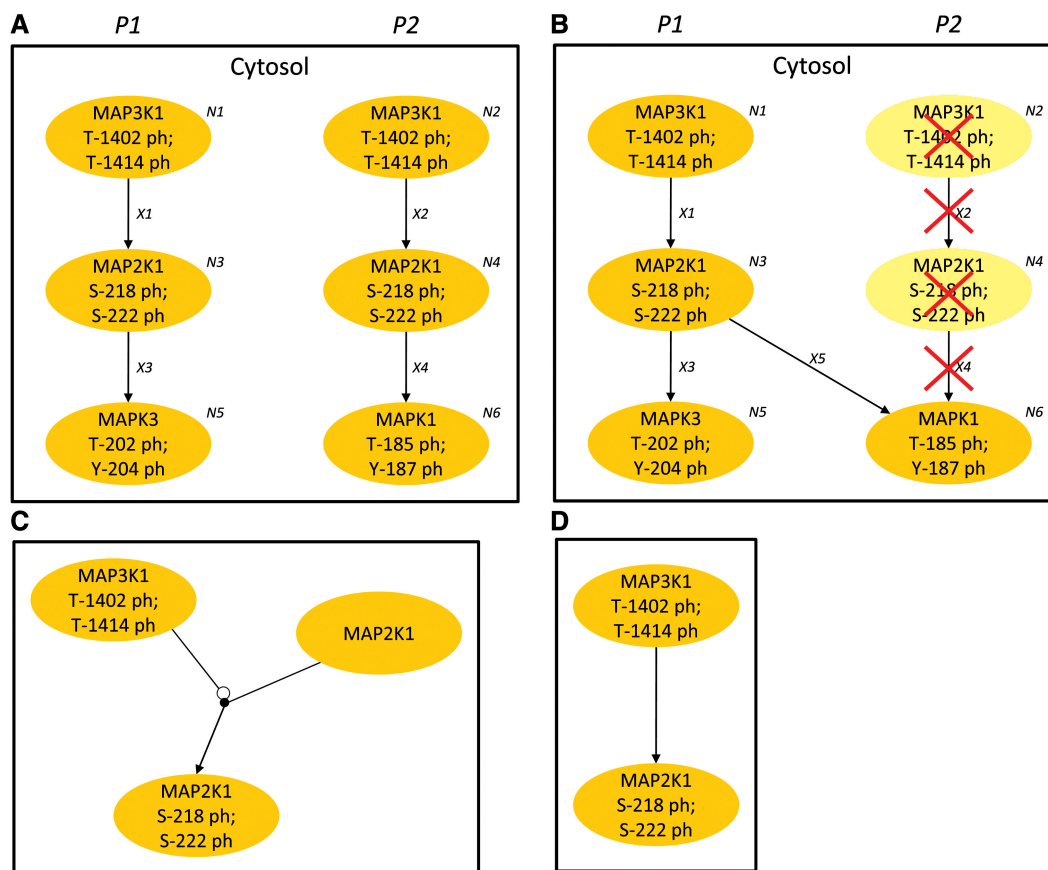
Figure 1. The PINT pathways-integration process. (**A**) and (**B**) are the general rules with which PINT performs pathways-integration process. (**C**) is a process diagram and (**D**) is an entity relationship diagram, where both are considered to be equivalent by PINT.

opposite direction, then in the merged model both edges will be preserved.

In the first rule, PINT trusts that a bi-directional edge represents a reversible reaction and assumes that its one-directional counterpart in other models is the result of insufficient annotation. In the second rule, PINT assumes that the two counterparts in opposite directions are likely to represent two distinct reactions, each catalyzed by an exclusive enzyme. Unfortunately, these rules for handling directionality may still result in errors in merged models if these assumptions are violated. Thus, users are advised to make sure that the private models they have uploaded contain well-curated information about the directionality of metabolic reactions.

Integrating pathways containing insufficient properties of entities. When a model annotated with insufficient compartment and modification information is integrated with a well-annotated model, the merged model of them may contain some redundant nodes that cannot be removed. Incorporating more annotations to an insufficient model can be a good solution, but may require considerably more efforts for literature review, model editing, etc.

Thus, PINT was also implemented with a tolerant integration mode, which may facilitate users to determine whether it is worth creating a better-annotated model to perform a more sensible BPI. The tolerant mode can

integrate pathway models in a less stringent way, which ignores entity compartment and modification information. In other words, if models are integrated in this way, the consistency of entity modification and/or compartment information will not be taken into consideration when determining the PIN sets as described above (Figure 1A and B and relevant text in Section The pathways integration rule). Although this operation may not generate a decently merged model, it can be used as a quick and dirty approach to instantly generate a vague but informative picture about how much crosstalk there might be among distinct pathways.

Integrating process-diagram models with entity-relationship-diagram models. Process diagram (PD) and entity relationship diagram (ERD) are two different notation styles to present pathways diagrammatically. Each of them has its unique advantages and attracts its own supporters owing to differential reasons (15). One concern is that a model specifically edited to be displayed in a PD style may contain temporal sequence of biological events (16), which an ERD compliant model usually lacks. Therefore, the aforementioned stringent rule used by PINT does not instantly suffice to integrate an ERD compliant model with a PD compliant one. Integrating only models exclusively in PD or ERD can be one quick and clean measure, at the cost of losing invaluable information in the models already excluded for one round of BPI.

In the following, we demonstrate the trick used by PINT to partially address this issue, without claiming that this PINT has provided a good solution. In the long run, we believe that a unified pathway notation style should be used instead, suggested by Le Novere *et al.* (15). Besides, this trick works only for diagrammatic preview in PINT, both types of notation will be preserved in the merged models exported from PINT, and users may further use SBML editors to resolve such inconsistency issues.

In an ERD compliant model, the state-transition information is usually excluded. For instance, in Figure 1D, the previous state of MAP2K1 is not found in this ERD, whereas such a piece of information is available in the PD as in Figure 1C. PINT uses the following steps to decide how to integrate such an ERD model with a PD model.

First, for each ERD model, PINT will take the following steps: (i) finding each edge that connects two nodes containing different entities; (ii) labeling each such edge as an 'undetermined modulation' (UM); (iii) labeling each node upstream/downstream of an UM as a functionally 'undetermined upstream/downstream reactant' (UUR/UDR). Second, for each PD model, PINT will look for the state transition information relevant to each UM, UUR and UDR.

An UUR may be a catalyst, an enhancer, an inhibitor or a subunit of a complex. An UDR usually corresponds to a product, whereas the reactant, namely the previous state of an UDR, is usually omitted in an ERD compliant model. For instance, in Figure 1D, 'MAP3K1 T-1402 ph; T-1414 ph' is an UUR; 'MAP2K1 S-218 ph; S-222 ph' is an UDR; the arrow between the UUR and UDR corresponds to an UM. When this ERD compliant model is integrated with the PD compliant model (Figure 1C), PINT can now decide that the UUR catalyzes the phosphorylation of the previous state of the UDR. In this case, PINT will consider that the two models are biologically equivalent and thus the merged model will be exactly the same as that shown in the PD (Figure 1C). Otherwise, if PD compliant models do not have the information that can compensate the insufficiency of an ERD compliant model, the relation of UM, UUR and UDR will be kept unchanged in the merged model.

The online annotation editor (for uploading SBML files only). In addition to the public pathway models, PINT users are allowed to upload and use their private models to perform pathways integration. However, such models are not necessarily annotated in a manner directly suitable for BPI. For instance, it is not uncommon that one protein is annotated with various synonyms in different models (e.g. both HER2 and ERBB2 correspond to the same protein); one generic name used in different models may actually refer to different proteins (e.g. BMP can be BMP2A, BMP2B, BMP3, *etc.*). The non-unique naming of entities in pathway models may make PINT fail to find correct inter-pathway linkers—the very first step in BPI. Besides, using curated models downloaded from public-domain databases does not necessarily generate an error-free merged model. For instance, in the curated

models provided by BioModels, phosphorylation of proteins across different models can be represented as suffixes to protein symbols in a variety of styles (e.g. '_p', '-PT', '-PY', '-p', etc.). Redundant nodes containing entities annotated in such inconsistent styles cannot be properly removed in BPI.

The PINT online annotation editor was hence designed to assist users to revise the inaccurate and/or insufficient entity annotations contained in an uploaded SBML model. The online editor tabulates the mapping between each node and its re-annotation suggested by PINT (Supplementary Figure S1A). The re-annotation is gene symbol-based, and is more intuitive for bench biologists to conceive than an accession number-based mapping. Internally, in this automatic re-annotation process, PINT maps each entity name that may be a synonym to its approved symbol based on the lookup tables downloaded from HGNC (17) and ChEBI (18). When there is one-to-many relationship between a synonym and a set of candidate symbols, PINT returns the most probable one with the aid of the internal 'frequently used synonym' versus 'most probable approved symbol' lookup table. We prepared this table through manually analyzing the human pathway models in BioModels (1).

Thus, each gene or gene product entity is re-annotated with a HGNC-approved gene symbol (17), its molecular type (e.g. DNA, RNA and protein), and its modifications if available (Supplementary Figure S1A). If an entity represents a small molecule such as water, ATP, GDP, etc., it will be re-annotated with a ChEBI-approved symbol (18). The case shown in Supplementary Figure S1A is about the PINT re-annotation of a node named Ras-GDP. In this node, there are two entities, a protein and a small molecule, which have been correctly re-annotated as HRAS (type: PROTEIN) and GDP (type: SIMPLE_MOLECULE) by PINT (Supplementary Figure S1A). Then users may accept the re-annotation or revise it again. To save users' efforts in repetitively updating the entity properties of multiple occurrences of one entity in a model, PINT provides a replace-all-identical-synonym function. When users click the 'Edit' button as shown in Supplementary Figure S1A, they will see a pop-up window like the one in Supplementary Figure S1B. Any changes of the properties made in this window can trigger the PINT online editor to replace the annotations of all its identical entities in this model being edited. On the other hand, if users do not want to perform a replace-all operation, they may instead choose to modify the properties of a particular occurrence of an entity in the textarea as shown in Supplementary Figure S1A.

Once the symbol of a protein entity in a human pathway model is changed by users, its annotation will also be updated with a new UniProt accession number corresponding to the new symbol. Although this internal property is invisible to users, it will be included in the SBML file of this model exported by PINT. This measure is to make a PINT-generated model provide information as suggested by MIRIAM (14).

Website usage

The PINT website provides an integrated environment for users to perform the functions that are required for BPI. The system flow is presented in Figure 2. On the PINT web interface, these functions are categorized into four web pages: Start, Upload, Browse and Workbench (see the menu bar items on the PINT web pages). Users may search for pathway models to start with, upload private pathway models or gene-symbol lists (for highlighting entities on merged diagrams), browse available models in PINT, and select pathway models to perform the core BPI functions, respectively. Uploaded models and gene-symbol lists will be cached in PINT temporarily and be cleaned every Saturday night.

The PINT public models and users' private models. PINT provides ~190 human pathway models for new users to try BPI immediately (Figure 2, Input layer, PINT public pathway models). These models were prepared based on two data sources. First, we used CellDesigner (16) to re-build ~30 human pathway models of the pathway

diagrams provided by BioCarta (http://www.biocarta.com/). Second, we retrieved ~160 human pathway models through parsing the human pathway information from the pathway XML file downloaded from PID (3).

Besides, users can upload their private models into PINT on the 'Upload' page. Pathway models to be uploaded can be either downloaded from BioModels (1) or Panther (2), or created by CellDesigner (version 3.5.1 or later) (16) or SBMLeditor (under implementation) (19). Besides, every PINT public pathway model can be downloaded in the SBML format, and then re-uploaded and re-annotated as users' private models. This operation is useful when integrating pathway models that contain fusion genes and their products, as demonstrated in an example on the PINT 'Documentation' page.

Selecting pathway models. Both the public and private pathway models already in PINT can be browsed or searched by users (Figure 2, Input layer). The PINT 'Browse' page lists alphabetically the public and private pathway models respectively. Besides, the PINT 'Start'
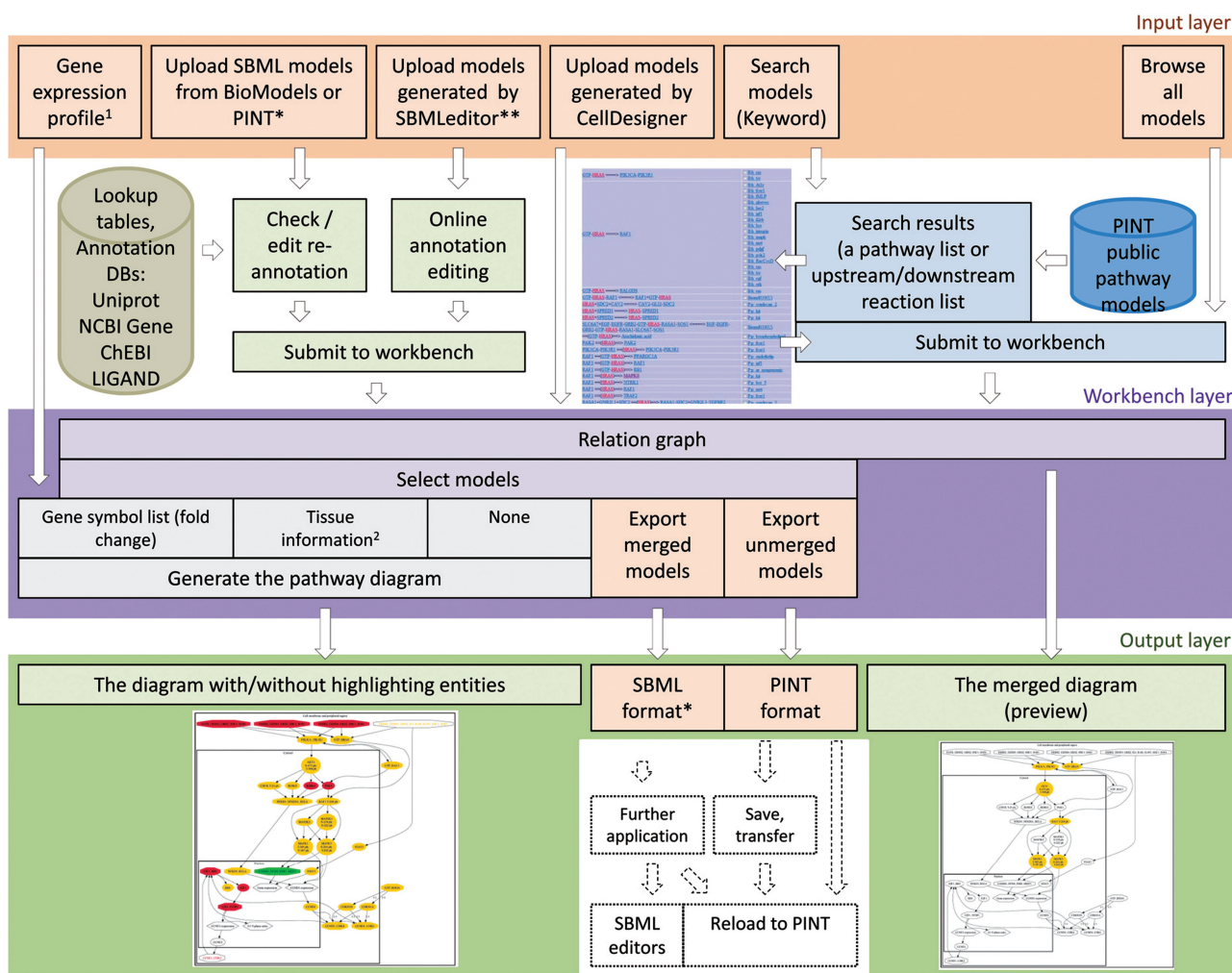
Figure 2. System flow of PINT. *Users can re-upload the SBML file exported from PINT. **PINT will accept the SBML files generated by SBMLeditor. [1]A gene list and fold change file should follow a particular format (Supplementary Figures S11 and S13). [2]Tissue information was retrieved from UniGene (Build 221) (http://www.ncbi.nlm.nih.gov/unigene).

page provides users with several search options. Users may type in keywords (e.g. gene symbol, synonym), or upload a long list of keywords. Users may also search for reactions involving a user-specified entity. This reaction-search function, very similar to the PID-style query (3), allows users to take an entity as the 'center point' to wander along relevant reactions further upward or downward, and thus enables users to quickly find upstream and downstream pathway models. From the result returned by PINT after a 'search' or 'browse' operation, users may choose and send pathway models to the PINT Workbench to perform BPI.

Integrating pathway models (PINT Workbench). The PINT Workbench (Figure 2, Workbench layer) provides the core BPI functions that allow users to visualize pairwise relations among pathway models, choose a subset of them for integration, visualize the merged diagram, highlight entities on the merged diagram and download the merged model for reuse in PINT or other SBML model editing/visualization tools (Figure 2, Output layer).

Relation graph. The core BPI functions of PINT can be activated if a number of pathway models have been submitted to the Workbench (see the previous section 'Selecting pathway models'). On the activated Workbench, there are a Relation graph, a list of pathway models, and options to display or export a merged model. The Relation graph shows the pairwise relations among the models listed in the PINT Workbench (for an example, see Supplementary Figure S6). Each number in this graph is to indicate how many nodes two potentially linkable pathway models have in common. When a number is clicked, PINT can generate a merged diagram of two relevant models for preview, where identical nodes commonly found in both models are shaded in yellow [Figure 2, Output layer, the merged diagram (preview)]. Users may thus decide whether such a merged model can be interesting or not, and choose relevant distinct models for integration at a later stage.

Highlighting entities on a merged diagram (Painting, PINT Workbench). The BPI conducted by PINT relies on putatively identical nodes (PIN) sets (see section 'The pathways integration rule'), but not on any gene or protein expression data. This means that merged models so generated are just hypothetical and may need further experimental evidence to support their biological significance. Therefore, PINT provides a 'painting' function that can highlight entities based on users' gene-symbol lists (for the sample format see Supplementary Figure S11), expression data or the tissue/histology information (Supplementary Figure S5, S8–S10) retrieved from UniGene (http://www.ncbi.nlm.nih.gov/unigene). This function can not only assist users to explore if their experimental data are consistent with a merged model but also help users to determine if a merged model makes biological sense. For instance, one approach to assess the biological role of a merged model is to highlight entities with the tissue/histology information. If on a merged pathway diagram there are more entities expressed in a cancer tissue (highlighted in red) than in the normal tissue (highlighted in green), perhaps the pathway represented by this merged model is implicated with this cancer. Thus, based on painted diagrams, bench biologists may use their expertise to decide whether the hypothetical pathway represented by a merged model is interesting, and therefore design more experiments to investigate its features. Alternatively, they may decide to try other combinations of models to perform another round of pathways integration and see if experimental evidence can support other merged models.

In addition to what is described above, the diagram painting function can be used to highlight entities in other biological context, such as selective pressure in evolution. To show this application, PINT provides the gene list containing the K_a/K_s values of human-mouse orthologous gene pairs (Supplementary Figure S12, '[single click] to upload K_a/K_s values...'). PINT can use these values to render nodes in the heatmap color scheme in pathway diagrams. Besides, users may also prepare their own gene lists that suit their research interests (for the sample format see Supplementary Figure S13).

An example case

Here, we present an example about how to use PINT to integrate three models of the RacCycD, Her2 and Ras pathways. The Ras pathway is about how Ras protein is involved in the inhibition of cell apoptosis (20). The Her2 pathway is about how Her2 protein is involved in the signaling related to cell growth and differentiation (21). The RacCycD pathway is about how Rac and CycD proteins are involved in the G1/S transition in the cell cycle (22). Since Her2 protein plays an important role in breast cancer (23), it is interesting to know how these three pathways may interact with each other to control cell growth and anti-apoptosis in breast cancer.

We used CellDesigner to re-build the three pathway models based on the diagrams provided by BioCarta (http://www.biocarta.com/) and then converted these three models into the SBML format. To generate a clear merged diagram for demonstration only, we have trimmed these models by taking out certain nodes. Besides, the entities in the 'cell membrane and peripheral region' compartment were moved into the 'Cytosol' compartment in order to reduce the number of compartments that will be displayed in a diagram. Users may download these models from NAR online (Supplementary data, RacCycd.xml, her2.xml, ras.xml). Figures S2–S10 in the Supplementary Data show the screenshots of the step-by-step procedures to upload these pathway models (Figures S2–S5), display the merged diagram for preview (Figures S6 and S7), highlight entities with their tissue/histology information on the merged diagram of the three pathway models (Figures S8 and S9), and list the detailed node information (Figure S10) for one node in this merged diagram.

This application not only eliminated the tedious manual drawing, but also generated a downloadable model for users. When this model is re-uploaded to PINT, all the annotations such as related reactions, external links, etc.

will be restored to the pathway diagram as well. This tool provides basic functions for both computational and experimental biologists to come up creative application scenarios.

DISCUSSION

PINT can integrate only pathway models that are SBML compatible, which must contain annotation about the role (reactant, product or catalyst) of each protein entity. Even though a complex can be an entity in PINT, this tool is not designed to integrate other biological networks in which the aforementioned annotations are unavailable. A typical example of such network is the high-throughput data of protein–protein interactions. Nevertheless, PINT is able to integrate different types of pathways from different organisms.

Although the pathways collected in PINT are human pathways, this tool may also integrate private pathway models from other organisms. As long as the entities are named consistently, PINT will be able to merge entities together. For example, EC number can be used as an entity name and can be merged correctly. The names of small molecules will be converted to the ChEBI names based on the alias tables from both KEGG and ChEBI. However, PINT cannot provide annotation to entities from other organisms at the present time.

To integrate the cross-species models, users need to have their own convention to name entities, such as the EC number. In this case, PINT may generate a reference map from the pathways in different species. Similarly, the pathways in different tissues of a given organism can be integrated. The 'painting' function can be used to label given entities on the reference map. These entities can be proteins from a given species, a given tissue or a given expression status at a given time point.

Taken together, PINT will be a useful tool not only in studying the regulatory circuit in a given cell/tissue types, but also in viewing the changes of pathway flux when dynamic information is available. Moreover, PINT can also be used to integrate models across states (e.g. different environments) or even across organisms (e.g. in a metagenomic context). In the latter case, the pathways across dependent species in different environments (24) can be explored.

FUTURE WORK

Currently, PINT does not support the integration of distinct kinetic equations that are contained in different models. We are working on finding useful guidelines to allow PINT to integrate such equations. Besides, the SBML models generated by PINT contain modification information that cannot be properly parsed by other SBML editors. We are discussing a revised format to resolve this issue.

SUMMARY

Since there are still many unknown details in biological networks, there will be updates to the current pathway models in the future. Biological pathway building and curation are an important step toward assembling a comprehensive picture of the networks. During this process, Biologists may need to try many combinations of available pathway models and to clarify if their experimental results fit a new merged model built on existing models. Therefore, biologists need flexible and user-friendly tools to assist them to go through the complex process and PINT was thus designed for this purpose.

AVAILABILITY

The web server of PINT is freely available at http://csb2 .ym.edu.tw/pint/.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Le Novere,N., Bornstein,B., Broicher,A., Courtot,M., Donizelli,M., Dharuri,H., Li,L., Sauro,H., Schilstra,M., Shapiro,B. *et al.* (2006) BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Res.*, **34**, D689–D691.
2. Mi,H. and Thomas,P. (2009) PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. *Methods Mol. Biol.*, **563**, 123–140.
3. Schaefer,C.F., Anthony,K., Krupa,S., Buchoff,J., Day,M., Hannay,T. and Buetow,K.H. (2009) PID: the Pathway Interaction Database. *Nucleic Acids Res.*, **37**, D674–D679.
4. Hucka,M., Finney,A., Sauro,H.M., Bolouri,H., Doyle,J.C., Kitano,H., Arkin,A.P., Bornstein,B.J., Bray,D., Cornish-Bowden,A. *et al.* (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, **19**, 524–531.
5. Viswanathan,G.A., Seto,J., Patil,S., Nudelman,G. and Sealfon,S.C. (2008) Getting started in biological pathway construction and analysis. *PLoS Comput. Biol.*, **4**, e16.
6. Babur,O., Colak,R., Demir,E. and Dogrusoz,U. (2008) PATIKAmad: putting microarray data into pathway context. *Proteomics*, **8**, 2196–2198.
7. Shannon,P., Markiel,A., Ozier,O., Baliga,N.S., Wang,J.T., Ramage,D., Amin,N., Schwikowski,B. and Ideker,T. (2003)

Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, **13**, 2498–2504.

8. Babur,O., Dogrusoz,U., Demir,E. and Sander,C. (2010) ChiBE: interactive visualization and manipulation of BioPAX pathway models. *Bioinformatics*, **26**, 429–431.

9. Hayashi,S. and Yamaguchi,Y. (2008) Estrogen signaling pathway and hormonal therapy. *Breast Cancer*, **15**, 256–261.

10. Lomonosova,E. and Chinnadurai,G. (2008) BH3-only proteins in apoptosis and beyond: an overview. *Oncogene*, **27(Suppl. 1)**, S2–S19.

11. Chonghaile,T.N. and Letai,A. (2008) Mimicking the BH3 domain to kill cancer cells. *Oncogene*, **27(Suppl. 1)**, S149–S157.

12. Jonsson,P.F. and Bates,P.A. (2006) Global topological features of cancer proteins in the human interactome. *Bioinformatics*, **22**, 2291–2297.

13. Jin,Q. and Esteva,F.J. (2008) Cross-talk between the ErbB/HER family and the type I insulin-like growth factor receptor signaling pathway in breast cancer. *J. Mammary Gland Biol. Neoplasia*, **13**, 485–498.

14. Le Novere,N., Finney,A., Hucka,M., Bhalla,U.S., Campagne,F., Collado-Vides,J., Crampin,E.J., Halstead,M., Klipp,E., Mendes,P. *et al.* (2005) Minimum information requested in the annotation of biochemical models (MIRIAM). *Nat. Biotechnol.*, **23**, 1509–1515.

15. Le Novere,N., Hucka,M., Mi,H., Moodie,S., Schreiber,F., Sorokin,A., Demir,E., Wegner,K., Aladjem,M.I., Wimalaratne,S.M. *et al.* (2009) The systems biology graphical notation. *Nat. Biotechnol.*, **27**, 735–741.

16. Kitano,H., Funahashi,A., Matsuoka,Y. and Oda,K. (2005) Using process diagrams for the graphical representation of biological networks. *Nat. Biotechnol.*, **23**, 961–966.

17. Bruford,E.A., Lush,M.J., Wright,M.W., Sneddon,T.P., Povey,S. and Birney,E. (2008) The HGNC Database in 2008: a resource for the human genome. *Nucleic Acids Res.*, **36**, D445–D448.

18. de Matos,P., Alcantara,R., Dekker,A., Ennis,M., Hastings,J., Haug,K., Spiteri,I., Turner,S. and Steinbeck,C. (2010) Chemical Entities of Biological Interest: an update. *Nucleic Acids Res.*, **38**, D249–D254.

19. Rodriguez,N., Donizelli,M. and Le Novere,N. (2007) SBMLeditor: effective creation of models in the Systems Biology Markup language (SBML). *BMC Bioinform.*, **8**, 79.

20. Pruitt,K. and Der,C.J. (2001) Ras and Rho regulation of the cell cycle and oncogenesis. *Cancer Lett.*, **171**, 1–10.

21. Casalini,P., Iorio,M.V., Galmozzi,E. and Menard,S. (2004) Role of HER receptors family in development and differentiation. *J. Cell. Physiol.*, **200**, 343–350.

22. Resnitzky,D., Gossen,M., Bujard,H. and Reed,S.I. (1994) Acceleration of the G1/S phase transition by expression of cyclins D1 and E with an inducible system. *Mol. Cell. Biol.*, **14**, 1669–1679.

23. Gullick,W.J. (1989) Expression of the c-erbB-2 proto-oncogene protein in human breast cancer. *Recent Results Cancer Res.*, **113**, 51–56.

24. Janga,S.C. and Babu,M.M. (2008) Network-based approaches for linking metabolism with environment. *Genome Biol.*, **9**, 239.